online commentary

Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract

Catherine ROBBE*, Calliope CAPON*†[1], Bernadette CODDEVILLE* and Jean-Claude MICHALSKI*

*Unité Mixte de Recherche CNRS/USTL 8576, Glycobiologie Structurale et Fonctionnelle, IFR 118, Université des Sciences et Technologies de Lille 1, 59655 Villeneuve d'Ascq Cedex, France, and †Laboratoire de Biochimie, Faculté des Sciences Pharmaceutiques et Biologiques de Lille 2, 59006 Lille Cedex, France

Purified human mucins from different parts of the intestinal tract (ileum, cecum, transverse and sigmoid colon and rectum) were isolated from two individuals with blood group ALe^b (A-Lewis^b). After alkaline borohydride treatment the released oligosaccharides were structurally characterized by nano-ESI Q-TOF MS/MS (electrospray ionization quadrupole time-of-flight tandem MS) without prior fractionation or derivatization. More than 100 different oligosaccharides, with up to ten monosaccharide residues, were identified using this technique. Oligosaccharides based on core 3 structures, GlcNAc(β1-3)GalNAc (where GlcNAc is *N*-acetyl-D-glucosamine and GalNAc is *N*-acetylgalactosamine), were widely distributed in human intestinal mucins. Core 5 structures, GalNAc(α1-3)GalNAc, were also recovered in all fractions. Moreover, a comparison of the oligosaccharide repertoire, with respect to size, diversity and expression of glycans and terminal epitopes, showed a high level of mucin-specific glycosylation: highly fucosylated glycans, found specifically in the small intestine, were mainly based on core 4 structures, GlcNAc-(β1-3)[GlcNAc(β1-6)]GalNAc, whereas the sulpho-Le^x determinant carrying core 2 glycans, Gal(β1-3)[GlcNAc(β1-6)]-GalNAc (where Gal is galactose), was recovered mainly in the distal colon. Blood group H and A antigenic determinants were present exclusively in the ileum and cecum, whereas blood group Sd^a/Cad related epitopes, GalNAc(β1-4)[NeuAc(α2-3)]Gal (where NeuAc is *N*-acetylneuraminate), were found to increase along the length of the colon. Our findings suggest that mucins create an enormous repertoire of potential binding sites for microorganisms that could explain the regio-specific colonization of bacteria in the human intestinal tract.

Key words: glycan, glycosylation, intestinal mucin, MS, mucin, O-linked oligosaccharide.

INTRODUCTION

Intestinal mucins, the major protein component of the mucus covering the epithelium of the gastrointestinal tract, are highly glycosylated macromolecules distinguished by the presence of dense O-glycosylation on the amino acids serine and threonine, clustered in 'mucin domains' [1–3]. Their functions include lubrication and modulation of water and electrolyte absorption [4]. In addition, mucins play an important role in protecting the underlying epithelium from mechanical and chemical stress. They may also have more specific activities, such as providing attachment sites for commensal and pathogenic microbes, and are also ligands for the targeting of leucocytes to endothelial cells [5–8]. Altered mucin glycosylation has been observed frequently and is associated with many pathologies, including cystic fibrosis, Crohn's disease and cancer [9,10]. Furthermore, different studies have indicated abnormal glycosylation in malignant transformation leading to shorter carbohydrate side chains or altered sialylation [11,12].

Our understanding of the role of intestinal mucin alterations in the disease state is limited by partial knowledge of the structural composition of normal mucin. Very few studies have been concerned with the glycosylation of normal human mucins in the whole colon [13,14] and in the descending colon [15], and little is known about the distribution of the various glycans in the different regions of the human intestine. Previously we demonstrated the presence of different gradients of oligosaccharides along the intestinal tract [16]. Using mainly NMR techniques, we showed the presence of decreasing gradients of fucose and ABH blood group expression from ileum to rectum and of an increasing acidic gradient along the gut.

In the present paper, we report the utilization of ESI Q-TOF MS/MS (electrospray ionization quadrupole time-of-flight tandem MS) to determine the detailed structure of oligosaccharide chains isolated from mucins from each part of the intestine (ileum, cecum, transverse and sigmoid colon and rectum). Mucin O-glycosylation has been studied mainly by purification of single oligosaccharides using HPLC followed by analysis with NMR and MS [17–19]. However, this general procedure is time consuming, requires the appropriate instrumentation and excessive amounts of glycoproteins. Accordingly, we adopted another approach based on MS, allowing a rapid profiling of O-glycan mixtures to provide sufficient structural information for the determination of differences in mucin glycosylation. Sequencing of oligosaccharides was acquired in both positive and negative ion modes to obtain complete elucidation of the structure for each mucin oligosaccharide. Positive ion mode is useful for the determination of the core-type oligosaccharide and the position of fucose, whereas negative ion mode produces characteristic cross-ring cleavages necessary for differentiation between isomers and identification of NeuAc (*N*-acetylneuraminate) and sulphate substitution [20].

The results show a remarkable degree of heterogeneity in human intestinal mucins, which is associated with extensive differences in the glycosylation of the various mucin populations between the different regions of the intestine.

Abbreviations used: ESI Q-TOF MS/MS, electrospray ionization quadrupole time-of-flight tandem MS; Fuc, fucose; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GalNAc-ol, *N*-acetylgalactosaminitol; GlcNAc, *N*-acetyl-D-glucosamine; HexNAc, *N*-acetylhexosamine; Le, Lewis; nano-ESI-MS/MS, nano electrospray ionization tandem MS; NeuAc, *N*-acetylneuraminate.

[1] To whom correspondence should be addressed (email calliope.capon@univ-lille1.fr).

© 2004 Biochemical Society

EXPERIMENTAL

Human samples and mucin preparation

All the immediate autopsy specimens were obtained from France Transplant Association from kidney donors according to protocols approved by the National Ethical Committee. Samples of mucosa were snapfrozen in liquid N_2 and stored in liquid N_2 until use. The human samples from which the different parts of the gut were analysed came from two male donors with ALe^b (A-Lewis^b) blood group and colon type W + Z+, two antigenic specificities designated W and Z by Zweibaum et al. [21] and described in normal colonic secretions. After thawing, the tissue was kept at 4 °C and the mucosa scraped, homogenized in distilled water (Ultra-Turrax, Jankee and Kunkel, Stauffer, Germany), and centrifuged (1 h at 48 000 ***g***). The supernatant was heated for 1 h in a boiling water bath and further centrifuged (1 h at 48 000 ***g***). The resultant supernatant was dialysed against distilled water for 2 days at 4 °C and freeze-dried [22].

Release of oligosaccharide alditols from mucin by alkaline borohydride treatment

The colonic mucins were submitted to β-elimination under reductive conditions (0.1 M KOH, 1 M KBH_4) for 24 h at 45 °C [23]. The mixture of oligosaccharide alditols was purified by size-exclusion chromatography on a column of Bio-Gel P2 [85 cm × 2 cm (internal diameter), 400 mesh; Bio-Rad, Richmond, CA, U.S.A.], equilibrated and eluted with water at 10 ml/h at room temperature (20 °C). The oligosaccharide fractions, detected at A_{206}, were pooled for structural analysis.

Fractionation of the oligosaccharide alditols by HPLC

The oligosaccharide alditols released from each part of the intestine were subjected to fractionation by HPLC (Dionex Chromeleon System, Sunnyvale, CA, U.S.A.) on a primary amino-bonded silica column (Supelcosyl, LC-NH_2, 4.6 mm × 250 mm; Supelco, Bellefonte, CA, U.S.A.) using a mixture of acetonitrile/30 mM H_2PO_4K/water (75:0:25, by vol.) to (50:50:0, by vol.) over 60 min with a flow rate of 1 ml/min. Oligosaccharides were detected by UV spectroscopy at 200 nm using an UVD 170U detector (Dionex).

Nano-ESI-MS/MS (nano electrospray ionization tandem MS)

All analyses were performed on a Q-STAR Pulsar quadrupole time-of-flight (Q-q-TOF) mass spectrometer (Applied Biosystems/MDS Sciex, Toronto, Canada) fitted with a nano-electrospray ion source (Protana, Odense, Denmark). Oligosaccharides dissolved in water (60 pmol/μl) were acidified by addition of an equal volume of methanol/0.1 % formic acid and sprayed from gold-coated 'medium-length' borosilicate capillaries (Protana). A potential of − 800 V was applied to the capillary tip and the focusing potential was set at − 100 V, the declustering potential varied between − 60 V and − 110 V. For the recording of conventional mass spectra, TOF data were acquired by accumulation of 10 MCA (multiple channel acquisition) scans over mass ranges of *m*/*z* 400–2000. In the collision-induced dissociation MS/MS analyses, multiple charged ions were fragmented using nitrogen as collision gas (5.3×10^{-5} torr, where 1 torr = 0.133 kPa), the collision energy varying between − 40 and − 90 eV to obtain optimal fragmentation. The collision-induced dissociation spectra were recorded on the orthogonal TOF analyser over a range of *m*/*z* 80–2000. Data acquisition was optimized to supply the highest possible resolution and the best signal-to-noise ratio, even in the case of low abundance signals. Typically, the FWHM (full width at half maximum) was 7000 in the measured mass ranges. External calibration was performed prior to each measure using a 4 pmol/μl solution of taurocholic acid in acetonitrile/water (50:50, v/v) containing 2 mM of ammonium acetate.

RESULTS

Mucins were isolated from different parts of normal human intestinal tract (ileum, cecum, transverse and sigmoid colon and rectum) of two blood group ALe^b individuals. Oligosaccharide alditols were released by base/borohydride treatment followed by ion-exchange chromatography and desalting. The percentage of saccharides that were recovered after release from the protein backbone was estimated based on monosaccharide compositional analyses of the whole mucin and of the released fraction. The yield was approx. 50 %, typical for this procedure. The relatively low yield of O-glycans is not likely to have selectively distorted the structural pattern, since re-release from preserved glycopeptides by a second β-elimination cycle gave the same pattern of glycans. Hence, in all, the glycans analysed are representative of the major mucins of the intestinal mucosa.

After the desalting step, oligosaccharide mixtures from each part of the intestine were analysed directly by ESI Q-TOF MS/MS, without prior derivatization or fractionation steps.

Structural characterization of underivatized mucin O-glycans was obtained by nano-ESI Q-TOF MS/MS in both positive and negative ion mode using a series of diagnostic ions [20]. These particular ions were useful for differentiation between isomeric molecules and provided information about core type, fucosylation, sialylation and/or sulphation.

Fragment annotations applied in this study were based upon the suggested nomenclature by Domon and Costello [24] and by Karlsson et al. [25]. The fragment ions obtained in this study were mainly A_i, B_i, Y_j, Z_j ions. Moreover, an α suffix was used to designate cleavages in the 6-linked branch and a β suffix for cleavage in the 3-linked branch from the GalNAc-ol (*N*-acetylgalactosaminitol).

Figure 1 represents the MS spectra acquired in negative ion mode, for the total oligosaccharide fractions from the five mucin populations of donor 1. Most of the ions could be related to $[M - H]^-$ ions of the theoretical oligosaccharide structures. The spectra illustrate a remarkable structural diversity: the true heterogeneity is likely to be even larger, since a substantial number of the molecular ions may represent oligosaccharide species with isomeric structures. To avoid confusion, the mass values used in the text, Tables and illustrated fragmentations are mainly nominal masses.

As shown in Tables 1 and 2 and in Figure 1, the ileum population contained the largest oligosaccharides, predominantly neutral and highly fucosylated (up to 3 residues of fucose were detected in these glycans). Among the 55 structures identified in this region, only 18 were acidic, whereas 35 were fucosylated. Moreover, the largest oligosaccharides were found exclusively in the ileum. These results could be explained in part by the fact that, in the other regions, glycans were mainly sialylated and/or sulphated, corresponding to chain termination.

Most of the oligosaccharides recovered in the cecum were also present in the ileum, except for some sulphated glycans. In the distal colonic mucins, glycans became more and more acidic and contained less fucose residues: in the ileum, approx. 30 % of oligosaccharides were acidic, whereas in the distal part more than 65 % contained sialic acid and/or sulphate residues.

© 2004 Biochemical Society

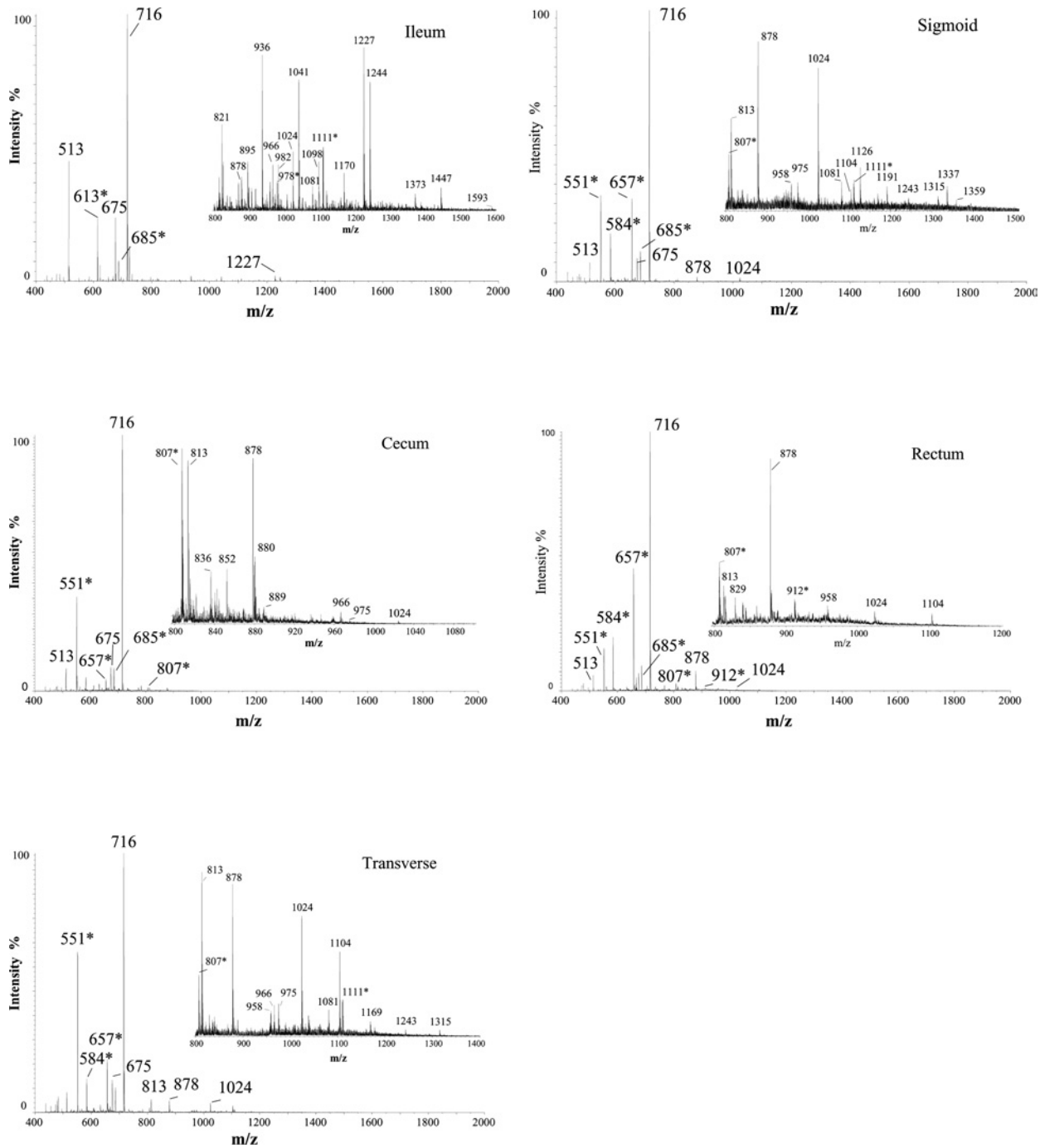

Figure 1 Mass spectra of the total oligosaccharides, acquired in the negative ion mode $[M-H]^-$, from ileum, cecum, transverse colon, sigmoid colon and rectum of donor 1

Signals marked with asterisks referr to dicharged ions $[M-2H]^{2-}$; ions at *m/z* 551, 584, 613, 657, 685 and 807 correspond respectively to the $[M-H]^-$ ions at *m/z* 1104, 1169, 1227, 1315, 1372 and 1615.

HPLC fractionation

Oligosaccharides released from each part of the intestinal tract were fractionated by HPLC on a primary amino-bounded column and the fractions collected from donor 1 sigmoid colon mucins were analysed by NMR spectroscopy and nano-ESI Q-TOF MS (results not shown). These studies not only allowed the confirmation of certain glycan structures, but above all were useful for the determination of the relative abundance of major oligosaccharides in this sample. Detection of glycans by UV at 200 nm is not as sensitive as MS and weakly expressed oligosaccharides cannot be identified. MS is not a quantitative method, since signal intensity is mainly dependent on the charge state and the size of the molecule. It was nevertheless interesting to note that the estimated abundance, deduced from peak intensity on MS spectra, of oligosaccharides having a similar charge density (i.e.

© 2004 Biochemical Society

Table 1 Proposed neutral oligosaccharide structures identified in human intestinal mucins by nano-ESI MS/MS

I, ileum; C, cecum; T, transverse colon; S, sigmoid colon; R, rectum; a, structures recovered from donor 2; b, structures recovered from donor 1. The upper branch of the oligosaccharides is indicated in bold.

		I		C		T		S		R	
Sequence/composition of oligosaccharide alditols	$[M+Na]^+$	a	b	a	b	a	b	a	b	a	b
GlcNAc → 3GalNAc-ol	449	+	+	+	+	+	+	+	+	+	+
GalNAc → 3GalNAc-ol	449	+	+	+	+	+	+	+	+	+	+
Gal → 3GlcNAc → 3GalNAc-ol	611	+	+	+	+	+	+	+	+	+	+
Gal → 3(**GlcNAc → 6**)GalNAc-ol	611	+	+	+	+	+	+	+	+	+	+
GlcNAc → 3Gal → 3GalNAc-ol	611	+	+	+	+	+	+	+	+	+	+
Fuc → 2Gal → 3GalNAc-ol	554	−	−	−	−	−	−	+	+	−	−
Gal → 3(Fuc → 4)GlcNAc → 3GalNAc-ol	757	+	+	+	+	+	+	+	+	+	+
Gal → 3(**Gal → 4GlcNAc → 6**)GalNAc-ol	773	+	+	+	+	−	−	+	+	−	+
Gal → 4GlcNAc → 3Gal → 3 → GalNAc-ol	773	+	+	+	+	−	−	−	−	−	−
HexNAc → Gal → 3GlcNAc → 3GalNAc-ol	814	+	+	+	+	+	+	+	+	+	+
HexNAc → 3Gal → 3(**GlcNAc → 6**)GalNAc-ol	814	+	+	+	+	+	+	+	+	+	+
Gal → 4GlcNAc → 3(**GlcNAc → 6**)GalNAc-ol	814	−	+	+	−	−	−	+	−	+	+
(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3GalNAc-ol	903	−	−	+	+	+	+	−	−	−	−
(Fuc →)GlcNAc → 3(Fuc → 2)Gal → 3GalNAc-ol	903	−	−	−	+	−	−	−	−	−	−
Gal → 4GlcNAc → 3(Fuc → 2)Gal → 3GalNAc-ol	919	+	+	−	−	−	−	−	−	−	−
(Fuc → 2)Gal → 3(**Gal → 4GlcNAc → 6**)GalNAc-ol	919	+	+	−	−	−	−	−	−	−	−
2 Gal, GlcNAc, Fuc, GalNAc-ol	919	+	+	+	+	+	−	−	−	+	+
HexNAc → 3Gal → 3(Fuc → 4)GlcNAc → 3GalNAc-ol	960	+	+	+	+	−	−	−	−	−	−
Gal, 2HexNAc, Fuc, GalNAc-ol	960	+	+	+	+	+	+	−	−	−	−
Gal, 2GlcNAc, Fuc, GalNAc-ol (core 2)	960	+	−	−	−	−	−	−	−	−	−
Gal, 2GlcNAc, Fuc, GalNAc-ol (core 4)	960	+	+	+	−	−	−	−	−	−	−
2 Gal, 2 HexNAc, GalNAc-ol	976	+	+	+	+	+	+	+	+	+	+
HexNAc → 3Gal → 4GlcNAc → 3(**GlcNAc → 6**)GalNAc-ol	1017	−	−	−	−	−	−	−	−	−	+
Gal → 3[(**Fuc → 2**)**Gal → 3**(**Fuc → 4**)**GlcNAc → 6**]GalNAc-ol	1065	+	−	−	−	−	−	−	−	−	−
(Fuc → 2)Gal → 3[(**Fuc → 2**)**Gal → 4GlcNAc → 6**]GalNAc-ol	1065	+	−	−	−	−	−	−	−	−	−
2 Gal, HexNAc, 2 Fuc, GalNAc-ol	1065	+	−	+	−	−	−	−	−	−	−
HexNAc → 3(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3GalNAc-ol	1106	+	+	+	+	−	−	−	−	−	−
2 Gal, 2 HexNAc, Fuc, GalNAc-ol	1122	+	+	+	+	−	−	−	−	−	+
Gal, 3 HexNAc, Fuc, GalNAc-ol	1163	−	−	−	−	−	−	−	−	−	+
HexNAc → 3Gal → 3(Fuc → 4)GlcNAc → 3(**GlcNAc → 6**)GalNAc-ol	1163	+	+	−	−	−	−	−	−	−	−
2 Gal, 3 HexNAc, GalNAc-ol	1179	−	−	−	−	−	−	+	+	+	+
2 Gal, HexNAc, 3 Fuc, GalNAc-ol	1211	+	−	−	−	−	−	−	−	−	−
2 Gal,2 HexNAc, 2 Fuc, GalNAc-ol	1268	+	+	+	+	+	+	−	−	−	−
(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3(**Gal → 4GlcNAc → 6**)GalNAc-ol	1268	+	+	−	−	−	−	−	−	−	−
3 Gal, 2 HexNAc, Fuc, GalNAc-ol	1284	+	+	−	−	+	+	−	−	−	−
HexNAc → 3(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3(**GlcNAc → 6**)GalNAc-ol	1309	+	+	−	−	−	−	−	−	−	−
2 Gal, 3 HexNAc, 1 Fuc, GalNAc-ol	1325	+	+	−	−	−	−	−	−	−	−
2 Gal, 2 HexNAc, 3 Fuc, GalNAc-ol	1414	+	+	−	−	−	−	−	−	−	−
3 Gal, 2 HexNAc, 2 Fuc, GalNAc-ol	1430	+	+	−	−	−	−	−	−	−	−
2 Gal, 3 HexNAc, 2 Fuc, GalNAc-ol	1471	+	+	−	−	−	−	−	−	−	−
2 Gal, 4 HexNAc, 1 Fuc, GalNAc-ol	1528	+	+	−	−	−	−	−	−	−	−
3 Gal, 2 HexNAc, 3 Fuc, GalNAc-ol	1576	+	+	−	−	−	−	−	−	−	−
2 Gal, 3 HexNAc, 3 Fuc, GalNAc-ol	1617	+	+	−	−	−	−	−	−	−	−
2 Gal, 4 HexNAc, 2 Fuc, GalNAc-ol	1674	+	+	−	−	−	−	−	−	−	−
3 Gal, 3 HexNAc, 3 Fuc, GalNAc-ol	1779	+	+	−	−	−	−	−	−	−	−
2 Gal, 4 HexNAc, 3 Fuc, GalNAc-ol	1820	−	+	−	−	−	−	−	−	−	−

ratio number of charged monosaccharides to the number of total monosaccharides), was very similar to results obtained from integration of peaks on HPLC profiles. Taken together, these results allowed us to estimate the relative proportions of the main structures detected in human colonic mucins (Table 3).

Core structures

As shown in Tables 1 and 2, the main core structure of mucins in each part of the intestine was core 3, GlcNAc(β1-3)GalNAc-ol (where GlcNAc is *N*-acetyl-D-glucosamine), as found previously [13–15]. More than half of the structures were built around the trisaccharide Gal(β1-3/4)GlcNAc(β1-3)GalNAc-ol (where Gal is galactose) extended with lactosamine chains, fucose, NeuAc and/or sulphate residues as described previously [26–28]. Several minor components were also recovered with core 1 structure Gal(β1-3)GalNAc-ol, substituted by GlcNAc, Fuc (fucose) or NeuAc. None of them was sulphated.

All MS/MS spectra acquired in positive ion mode of heavy structures, e.g the highly fucosylated chains in the ileum, presented diagnostic ions of core 4 oligosaccharides, GlcNAc(β1-3)-[GlcNAc(β1-6)]GalNAc-ol, with ions at *m*/*z* 413, 431, 449 and 634, 652. None of these structures was sialylated or sulphated. Figure 2 represents the ESI-MS spectrum of a core 4 glycan at *m*/*z* 1163. Due to the core 4 symmetry it was difficult to obtain sufficient information, under the conditions used, to determine whether the fragment ions observed could be attributed either to the lower or the upper branch. Interpretations of the fragmentation for the lower branch are reported in the Figure 2.

© 2004 Biochemical Society

Table 2 Proposed acidic oligosaccharide structures identified in human intestinal mucins by nano-ESI MS/MS

I, ileum; C, cecum; T, transverse colon; S, sigmoid colon; R, rectum; a, structures recovered from donor 2; b, structures recovered from donor 1. The upper branch of the oligosaccharides is indicated in bold.

Sequence/composition of oligosaccharide alditols	$[M-H]^-$	I		C		T		S		R	
		a	b	a	b	a	b	a	b	a	b
Oligosaccharides with one NeuAc residue											
NeuAc → 6 GalNAc-ol	513	+	+	+	+	+	+	+	+	+	+
Gal → 3(**NeuAc → 6**)GalNAc-ol	675	+	+	+	+	+	+	−	−	+	+
(NeuAc → 3)Gal → 3GalNAc-ol	675	−	−	−	−	−	−	+	+	−	−
GalNAc → 3(**NeuAc → 6**)GalNAc-ol	716	+	+	+	+	+	+	+	+	+	+
GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	716	+	+	+	+	+	+	+	+	+	+
(Fuc → 2)Gal → 3(**NeuAc → 6**)GalNAc-ol	821	+	−	+	−	−	−	−	−	−	−
Gal → 3GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	878	+	+	+	+	+	+	+	+	+	+
(NeuAc → 3)Gal → 3(**GlcNAc → 6**)GalNAc-ol	878	−	−	−	−	−	−	−	+	−	−
(NeuAc → 3)Gal → 4GlcNAc → 3GalNAc-ol	878	−	−	−	−	−	−	−	+	−	−
GlcNAc → 3Gal → 3(**NeuAc → 6**)GalNAc-ol	878	+	−	+	−	+	+	+	+	+	+
Gal → 3(Fuc → 4)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1024	+	+	+	+	+	+	+	+	+	+
Gal → 4(Fuc → 3)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1024	+	+	+	+	+	+	+	+	+	+
(NeuAc → 3)Gal → 4(Fuc → 3)GlcNAc → 3GalNAc-ol	1024	−	−	−	−	−	−	−	+	−	−
HexNAc → 3Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1081	−	+	−	+	−	−	−	−	−	−
GalNAc → 4(NeuAc → 3)Gal → 4GlcNAc → 3GalNAc-ol	1081	−	−	−	−	+	+	+	+	+	+
GalNAc → 4(NeuAc → 3)Gal → 3GlcNAc → 3GalNAc-ol	1081	−	−	−	−	+	+	+	+	+	+
(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1170	+	+	+	+	+	−	−	−	−	−
HexNAc → 3Gal → 3(Fuc → 4)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1227	+	+	+	+	+	−	−	−	−	−
HexNAc → 3(Fuc → 2)Gal → 3GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1227	+	−	+	−	−	−	−	−	−	−
(NeuAc → 3)Gal → 4GlcNAc → 3Gal → GlcNAc → 3GalNAc-ol	1243	−	−	−	−	−	−	+	+	−	−
HexNAc → 3(Fuc → 2)Gal → 3(Fuc → 4)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1373	+	+	+	+	−	−	−	−	−	−
NeuAc → 3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 3[**Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	1697	−	−	−	−	−	−	+	+	+	+
Oligosaccharides wih one sulphate residue											
(SO_3^-)3Gal → 4GlcNAc → 3GalNAc-ol	667	−	−	−	+	−	−	+	+	+	+
Gal → 4(SO_3^-)6GlcNAc → 3GalNAc-ol	667	+	−	−	+	−	−	+	+	+	+
(Fuc → 2)Gal → 4(SO_3^-)6GlcNAc → 3GalNAc-ol	813	+	−	−	+	−	−	−	−	−	+
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3GalNAc-ol	813	−	−	+	+	+	+	+	+	+	+
Gal → 3[(**SO_3^-)3Gal → 4GlcNAc → 6**]GalNAc-ol	829	−	−	−	−	+	−	+	+	+	+
Gal → 3[**Gal → 4(**SO_3^-)**6GlcNAc → 6**]GalNAc-ol	829	−	−	−	−	+	−	−	−	+	+
(SO_3^-)3Gal → 4GlcNAc → 3Gal → 3GalNAc-ol	829	−	−	+	+	+	−	−	−	−	−
Gal → 3[(**SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	975	−	−	−	+	+	+	−	+	+	+
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 3GalNAc-ol	975	−	−	−	−	+	−	+	+	+	+
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 3[**Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	1486	−	−	−	−	−	−	−	+	−	−
Oligosaccharides with two acidic residues											
(SO_3^-)3Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	958	−	−	−	−	+	+	+	+	+	+
(NeuAc → 3)Gal → 3(**NeuAc → 6**)GalNAc-ol	966	+	+	+	+	+	+	+	+	−	−
(SO_3^-)3Gal → 3[(**SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	1055	−	−	−	−	−	−	−	+	−	−
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1104	+	+	+	+	+	+	+	+	+	+
(NeuAc → 3)Gal → 3[(**SO_3^-)3Gal → 4GlcNAc → 6**]GalNAc-ol	1120	−	−	−	−	−	−	+	−	+	+
(NeuAc → 3)Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1169	+	+	−	+	+	+	+	+	+	+
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 3(**NeuAc → 6**)GalNAc-ol	1266	−	−	−	−	+	+	−	−	−	−
(NeuAc → 3)Gal → 4(Fuc → 3)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1315	−	−	+	+	+	+	+	+	+	+
(SO_3^-)3Gal → 4GlcNAc → 3Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1323	−	−	−	+	−	−	−	+	−	−
GalNAc → 4(NeuAc → 3)Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1372	−	−	−	+	+	+	+	+	+	+
GalNAc → 4(NeuAc → 3)Gal → 3GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1372	−	−	−	+	+	+	+	+	+	+
(SO_3^-)3Gal → 4GlcNAc → 3Gal → 3[(**SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	1420	−	−	−	−	−	−	−	+	−	−
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 4GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1469	−	−	+	+	−	−	+	+	+	−
3 Gal, 2 HexNac, NeuAc, SO_3^-, GalNAc-ol	1485	−	−	−	−	−	−	−	−	+	−
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 3[(**SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 6**]GalNAc-ol	1566	−	−	−	+	−	−	+	+	+	+
(SO_3^-)3Gal → 4(Fuc → 3)GlcNAc → 3Gal → 4(Fuc → 3)GlcNAc → 3(**NeuAc → 6**)GalNAc-ol	1615	−	−	+	+	+	+	+	+	+	+
2 Gal, 2 HexNAc, Fuc, 2 NeuAc, GalNAc-ol	1680	−	−	+	−	−	−	−	−	−	−
2Gal, 2 HexNAc, 2 Fuc, 2 NeuAc, GalNAc-ol	1826	−	−	−	−	−	−	−	−	+	−

In the distal colon (sigmoid and rectum), core 2 structures Gal(β1-3)[GlcNAc(β1-6)]GalNAc-ol, were found to be mainly sulphated. Figure 3 shows the MS/MS spectrum in negative ion mode of the dicharged ion $[M-2H]^{2-}$ at *m/z* 527, corresponding to a core 2 oligosaccharide at *m/z* 1055 carrying two sulphate residues. The core was easily identified on the fragmentation spectrum by the presence of the two Y_1 ions: the $Y_{1\beta}$ fragment ion at *m/z* 813 indicated that a branch consisting of Gal, GlcNAc, Fuc and a sulphate group was linked to the GalNAc-ol, whereas the $Y_{1\alpha}$ fragment ion at *m/z* 464 indicated that another branch comprised one residue of Gal with a sulphate group linked to the GalNAc-ol.

One structure based on core 5, GalNAc(α1-3)GalNAc-ol (where GalNAc is *N*-acetylgalactosamine), was found in mucins

© 2004 Biochemical Society

Table 3 Estimated percentage of major oligosaccharides recovered by HPLC

I, ileum; C, cecum; T, transverse colon; S, sigmoid colon; R, rectum; n.d., not detectable.

	Oligosaccharide recovered (%)				
*m/z**	I	C	T	S	R
425†	n.d.	n.d.	n.d.	5.6	n.d.
513	3	2	2	2	2
551‡	n.d.	6	8	6	4
584‡	n.d.	5	1.5	3	2.5
587	9	20	11	9	5
613‡	12	–	–	–	–
657‡	–	2	3.5	3	2.5
675	3	2	3	2	3
685‡	–	5.5	13.5	13	7
716§	24	6	13	9	4
716‖	23	22.5	22	15	7
813	2	2	4	11	24
878	4	6	7	6	3
1024	n.d.	3	8	3	4
1081	–	5	6	8	4.5

* The HPLC fractions were identified by their corresponding *m/z* ions to allow comparison with the mass spectra.

† Ion at *m/z* 425 corresponding to both neutral core 3 and core 5 structures.

‡ Dicharged ions $[M-2H]^{2-}$.

§ Ion at *m/z* 716 corresponding to a sialylated core 5.

‖ Ion at *m/z* 716 corresponding to a sialylated core 3.

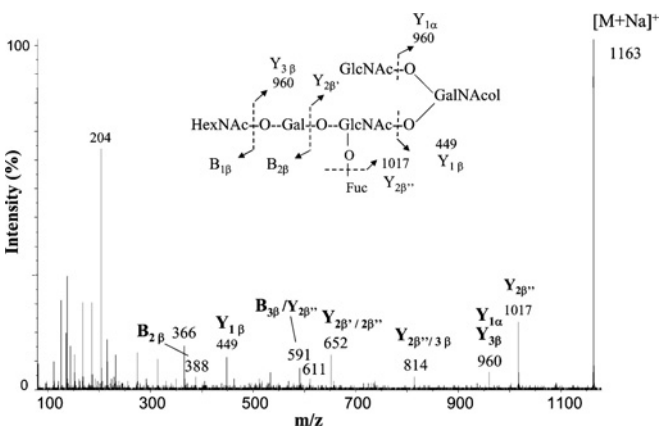

Figure 2 MS/MS spectrum of a core 4 oligosaccharide at *m/z* 1163, recorded in the positive ion mode $[M+Na]^+$

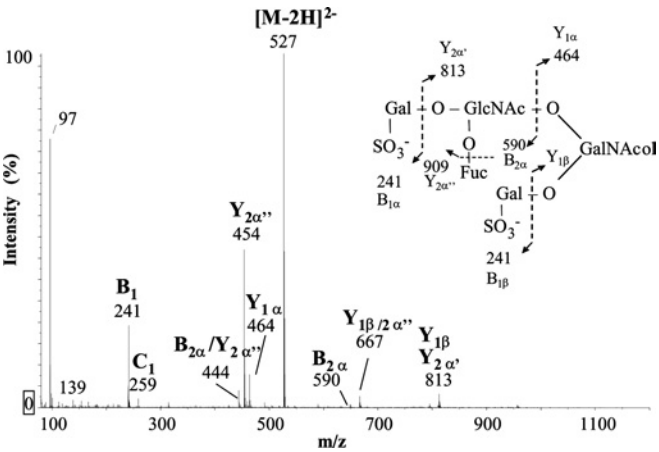

Figure 3 MS/MS spectrum, recorded in the negative ion mode, of the dicharged ion $[M-2H]^{2-}$ corresponding to a core 2 oligosaccharide at *m/z* 1055 with two sulphate residues

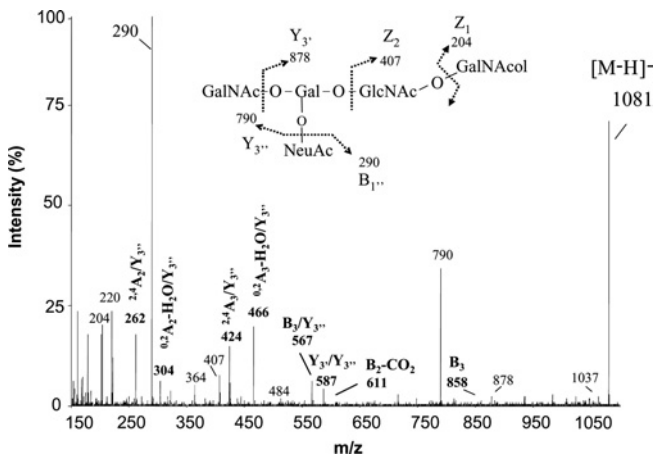

Figure 4 MS/MS spectrum of the $[M-H]^-$ ion at *m/z* 1081 corresponding to an oligosaccharide carrying a blood group Sd^a/Cad antigen, isolated from human sigmoid mucins

from the ileum to the rectum. This core was detected by NMR analysis on the mixture of oligosaccharides [16], since MS does not allow discrimination between GalNAc and GlcNAc residues.

Sd^a/Cad-related determinants

Two ions at *m/z* 1081 and 1372 respectively were identified on ESI-MS spectra of deprotonated oligosaccharides from colonic mucins, corresponding to structures with blood group Sd^a/Cad-like determinant, GalNAc(β1-4)[NeuAc(α2-3)]Gal. These ions were not recovered in the ESI-MS spectrum of ileal mucins.

For example, the MS/MS spectrum of the ion at *m/z* 1081 is shown in Figure 4. The major fragment ions were due to the loss of sialic acid (*m/z* 790) and to the $B_{1''}$ fragment ion at *m/z* 290 (a residue of sialic acid), but additional glycosidic cleavages producing mainly Y_i fragment ions enabled reconstruction of the glycan sequence. The Y_2 ion at *m/z* 425 (or Z_2 ion at *m/z* 407) indicated a GlcNAc residue linked to GalNAc-ol, giving a core 3 structure, and the $Y_{3'/3''}$ ion at *m/z* 587 indicated the addition of a Gal residue to this structure. No diagnostic ion at *m/z* 513 was detected in the spectrum, suggesting that the NeuAc residue was α2-3-linked to the Gal residue [20]. This last item of data constituted the proof that non-reducing terminal HexNAc (*N*-acetylhexosamine) was in fact a GalNAc. Indeed, the only structure recovered in human mucins with a terminal HexNAc linked to a sialylated Gal is a structure with the blood group Sd^a/Cad determinant. The addition of this terminal GalNAc(β1-4) linked to the Gal residue was observed from two fragment ions at *m/z* 304 and 790 respectively. The ion at *m/z* 304, corresponding to a $^{0,2}A_2$-H2O/$Y_{3''}$ fragment ion, was the diagnostic ion for the β1-4 linkage between GalNAc and Gal residues, and the ion at *m/z* 466 ($^{0,2}A_3$-H2O/$Y_{3''}$) confirmed the β1-4 linkage between Gal and GlcNAc [29].

© 2004 Biochemical Society

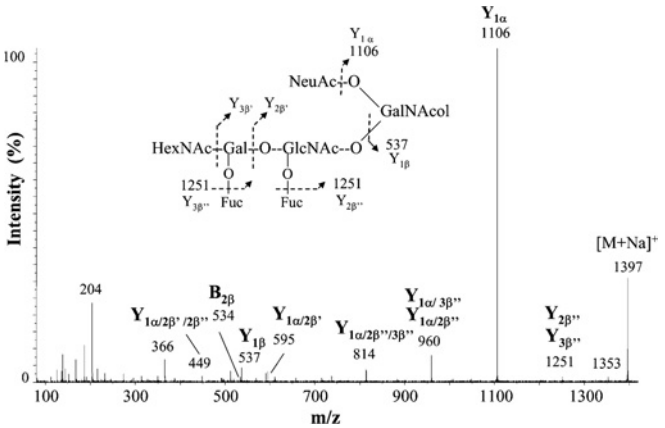

Figure 5 MS/MS spectrum of the $[M + Na]^+$ ion at *m/z* 1397 corresponding to an O-glycan with a blood group ALe^b determinant

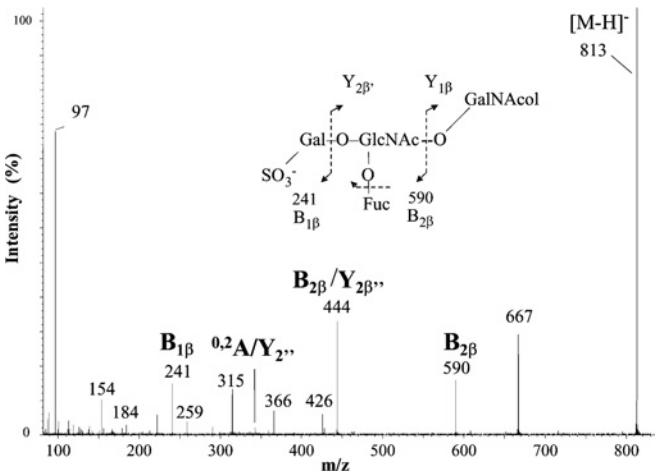

Figure 6 MS/MS spectrum of the $[M - H]^-$ ion at *m/z* 813 corresponding to an O-glycan carrying a Le^x determinant

The MS/MS spectrum of ion at *m/z* 1372 (results not shown) was similar to the first spectrum at *m/z* 1081, except for the presence of a fragment ion at *m/z* 513, diagnostic of a NeuAcα2-6 linked to the GalNAc-ol, giving a sialylated core 3 structure: GalNAc(β1-4) [NeuAc(α2-3)] Galβ1-3/4GlcNAc(β1-3)[NeuAc-(α2-6)]GalNAc-ol [20].

In a previous paper [16], we have shown that NMR analysis allowed us to confirm the presence of NeuAc(α2-3) with the characteristics of Sd^a/Cad determinants in colonic mucins from donor 1 only. Using nano-ESI Q-TOF MS/MS, it was possible to identify unambiguously the presence of these epitopes in donor 2. The two ions at *m/z* 1081 and 1372 were detected in the two spectra, but these ions were less intense in the ESI-MS spectrum of donor 2. Thus these results confirmed the previous NMR analysis: it seems that Sd^a/Cad determinants were under-expressed in donor 2 (results not shown).

Blood group ABH determinants

Several ions were only found in the ESI-MS spectra of ileum and cecum mucins, at *m/z*, for example, 1106, 1309 or 1397 in the positive mode. MS/MS fragmentation of the ion at *m/z* 1397 indicated the presence of an Le^b determinant, Fuc(α1-2)Gal(β1-3)[Fuc(α1-4)]GlcNAc, linked to the GalNAc-ol (Figure 5). Interpretations of the fragmentation are given in the Figure. The spectrum also indicated the presence of an additional HexNAc at the non-reducing end of the carbohydrate chain. ESI-MS does not discriminate between terminal GalNAc or GlcNAc residues (except in the case of an Sd^a/Cad determinant), but NMR analysis of the mixture of oligosaccharides for these two mucins (ileum and cecum) has shown the presence of blood group A determinants [16]. Thus the structure of this ion at *m/z* 1106 could be deduced as carrying the blood group ALe^b. Other similar structures were recovered in the first two regions of the intestine. Terminal HexNAc residues could be identified as GalNAc by similar arguments as given above, but it is impossible to confirm unambiguously the presence of blood group A epitopes by MS alone.

Acidic glycans

Most of oligosaccharides in colonic mucins carried acidic components with NeuAc and/or sulphate residues. These components were easily identified in the ESI-MS/MS spectra by their corresponding fragment ion at *m/z* 290 for the NeuAc residue and 97 for the sulphate residue [20].

NeuAc residues were mainly α2-6 linked to the GalNAc-ol ($Y_{1\beta}$ fragment ion at *m/z* 513 in the negative ion mode). The NeuAc(α2-3)Gal sequence was also found in the distal colonic mucins. Sialyl Tn, sialyl T and the sialylated core 3 structure GlcNAc(β1-3)[NeuAc(α2-6)]GalNAc-ol were recovered in all fractions.

The location of the sulphate group to Gal or GlcNAc residues was distinguished by the presence of ions at either *m/z* 241 [Gal-SO_3^-]$^-$ or *m/z* 282 [GlcNAc-SO_3^-]$^-$ on the spectra. Sulphate was essentially 3-linked to Gal or 6-linked to GlcNAc residues as described previously [16], and sulphated glycans were mainly recovered in distal colonic mucins (sigmoid colon and rectum regions).

Fucosylated structures

Fucose was present in a large variety of terminal linkages (Tables 1 and 2) that included blood group H epitopes [Fuc(α1-2)Gal(β1-3)GlcNAc], as well as Le^a {Gal(β1-3)[Fuc(α1-4)]-GlcNAc}, Le^b {Fuc(α1-2)Gal(β1-3)[Fuc(α1-4)]GlcNAc} and Le^x {Galβ(1-4)[Fuc(α1-3)]GlcNAc} determinants. High mass, highly fucosylated oligosaccharides, carrying the Le^b epitope, were mainly recovered in small intestine and cecum regions, whereas other components, such as sulpho-Le^x determinants, were found exclusively in colonic mucins. For example, in the MS/MS fragmentation spectrum in negative ion mode of the ion at *m/z* 813, originating from the sigmoid colon segment, a series of Bi fragment ions at *m/z* 241, 444 and 590 respectively indicated the sequence of the oligosaccharide (Figure 6). Moreover, cross-ring fragment ions $^{0,2}A/Y_{2''}$ at *m/z* 361 and 343 (loss of water) were diagnostic of a β1-4 linkage between Gal and GlcNAc residues, confirming the presence of an Le^x determinant on this glycan.

DISCUSSION

The mucins synthesized and secreted by the mucosal cells lining the gastrointestinal tract form a vital part of the total mucosal secretion giving rise to the extracellular barrier. The supramucosal barrier forms the primary line of defence against attack

© 2004 Biochemical Society

from the external environment, including micro-organisms, biochemical agents and mechanical stresses. *In situ* hybridization has demonstrated that mucins from each section of the gastrointestinal tract have a typical *MUC* gene composition: MUC2 is localized to goblet cell vesicles in both small and large intestine, whereas MUC3, more abundant in the small intestine, is found in most surface epithelial cells and in the upper parts of the crypt cells. Moreover, MUC5B and MUC6, weakly expressed in the colon, are absent in other parts of the digestive tract, whereas MUC4 is widely expressed in the gastrointestinal tract [30–33]. In the present paper, we report a detailed structural analysis of O-glycans by physical methods, providing evidence of a regio-specific glycosylation of mucins along the intestinal tract. Although many structures were common to all of the mucin species, each was noted to contain structures not found in the other mucin fractions. The pattern of *MUC* genes for each part of the intestine was not determined in this study; nevertheless, it can be hypothesized that differences in glycosylation may be correlated with *MUC* gene expression, suggesting that each mucin core protein may have its own glycosylation and that mucins may be distinguished structurally as well as functionally.

Using NMR spectroscopy and MALDI-MS techniques, previous investigations allowed us to demonstrate the presence of different gradients of oligosaccharides along the intestinal tract, with, for example, the establishment of an acidic gradient from the ileum to the rectum, associated with a decreasing gradient of Fuc [16]. In the present study, nano-ESI Q-TOF MS/MS was used for the structural characterization of mucin oligosaccharides from each part of the intestine, i.e. the ileum, the cecum, the transverse colon, the sigmoid colon and the rectum. The results are in agreement with the previously published colonic mucin analysis [13–15], in that almost all the structures are based on core 3. However, we also demonstrated structures with core 1 and 5 in each mucin. Core 5 was described previously by Kurosaka et al. [34] in a human rectal adenocarcinoma glycoprotein and by Feeney et al. [35] in human meconium, and was considered as a tumour marker. Our results suggest that core 5 glycans are structures recovered in normal human intestine, and are not diagnostic of a pathological situation.

More interestingly our observations show the presence of highly fucosylated core 4 oligosaccharides in the ileum and the identification of core 2 structures, carrying mainly sulpho-Le^x determinants in the distal colon (sigmoid and rectum). These structures were not recovered in previous studies [15]. The differences between these two reports may be attributed to the present use of nano-ESI Q-TOF MS, which is known to be more sensitive than NMR spectroscopy, enabling the detection and sequencing of minor components. Nevertheless, differences in the abundance of these core structures along the gut may imply evidence for different pathways in the glycosylation process of human intestinal mucins.

Differences in glycosylation between the five mucin populations affect not only core glycans, but also length and terminal epitopes. For example, blood group ALe^b determinants are recovered only in the small intestine and in the cecum, whereas blood group Sd^a/Cad-like structures are found largely in the distal colonic mucins. This last result is in agreement with a recent study showing a strong expression of the gene coding for the β1-4GalNAc transferase in distal colon [36]. The chain length of the mucin O-glycans from the ileum is demonstrated to be much longer and to carry mainly fucosylated terminal epitopes, such as blood group H or Le^b, whereas oligosaccharides from distal colonic mucins carry mainly NeuAc or sulphate residues. These differences in the fucosylation or sialylation of the five mucin populations may reflect an important mechanism in the regulation of mucin glycosylation. While the glycosylation was demonstrated to vary between each intestinal region, it is tempting to speculate about the preserved glycosylation between individuals, since our preceding paper [16] showed similar gradients of oligosaccharides for the two individuals studied, and results obtained herein by ESI-MS/MS analysis show no differences in glycan composition and structures between these two donors.

The presence of Gal1-3GlcNAc or Gal1-4GlcNAc appears to be selective for some locations and some peripheral acidic additions in particular sulphation. In our study, sulphates seem to be largely on Gal1-4GlcNAc units. It has been shown previously that the presence of sulphate on the C-3 of Gal in the Gal(β1-4)GlcNAc moieties increased the efficiency of these acceptors by 4- to 5-fold for Le α1-3 fucosyltransferases [37], leading to a structural epitope with the potential for recognition by specific receptors or lectins. This may suggest a further selective enzymic synthesis of identified structures in specific locations in the intestine, since sulphation increases from ileum to rectum.

Nano-ESI Q-TOF MS/MS allows the identification of more than a hundred different glycan structures in human intestinal mucins. This number is probably underestimated, because it is difficult to distinguish between isomeric structures by MS. This is important as the complexity and number of potential isomers grow quickly as a function of the number of sugar residues. It is likely that intestinal mucins carry several hundred different oligosaccharides. Part of this vast structural diversity could originate from bacterial degradation of mucins, since the intestinal microflora is believed to digest mucin oligosaccharides by secreting various linkage specific exoglycosidases [38–40].

In addition to the potential to bind lectins, it has been appreciated for some time that mucosal oligosaccharides are implicated as ligands for bacterial adhesion [41]. Therefore it can be surmised that one function for the diversity of oligosaccharide chains is to allow many interactions with micro-organisms which, in the gastrointestinal tract, for example, may help to maintain the rich bacterial flora. Together, the results of our recent work show a high degree of diversity in expression of glycans from each part of the intestine. The end result is an enormous repertoire of potential binding sites for micro-organisms that could explain the regio-specific colonization of bacteria in the human intestine.

The presence of large numbers of micro-organisms is usually associated with the secretion of sulphated mucin species. It has been proposed that sulphated mucins are particularly involved in protection against bacterial attack, because sulphation confers resistance to enzymic degradation of the mucus barrier by bacterial glycosidases or by host proteases, e.g. pepsin [42,43]. Also sialic acid residues have been implicated in the protease inhibition by mucins [44,45]. Microbiological studies have demonstrated high bacterial concentrations in the distal colon. These facts could explain in part the increasing number of acidic structures recovered along the intestinal tract.

Information about mucin glycosylation in each region of the intestinal tract will provide a comparative base to determine whether alterations in the oligosaccharide repertoire are associated with particular disease states. Many studies suggest that mucin glycan structures in cancer cells differ from those in normal cells [9,10,46]. Mucins from patients with intestinal cancer have less total carbohydrate and less sulphation, and the O-glycans are truncated and sialylated. Structural and antigenic properties of glycoproteins are changed with an increasing expression of Tn, sialyl Tn, sialyl T and Le antigens. The activity of the GlcNAc transferase responsible for the core 3 synthesis decreases

© 2004 Biochemical Society

in colon cancer, together with an increased activity of core 1 and 2 transferases [9]. In conclusion, the few structural studies of mucin glycans from colon cancer suggest that glycosylation is often altered in such pathologies. The results reported in the present work give us the necessary background information to begin studies to determine whether individuals who express certain carbohydrate epitopes on specific mucins are predisposed to diseases, such as cancer or Crohn's disease.

In conclusion, intestinal mucus is composed of several mucins that differ in size, glycosylation and tissue distribution. Differential secretion of mucins from the surface epithelium and glands may provide a mechanism for modulation of the composition of the protective mucus layer related to acid secretion or the presence of bacteria and noxious agents in the lumen.

We are grateful to Guy Ricart for his skillful technical assistance with the MS analysis and to Philippe Delannoy for careful re-reading of the manuscript. This investigation was supported in part by the CNRS (Unité Mixte de Recherche CNRS/USTL 8576; Director, J.-C. M.), by the Ministère de la Recherche et de l'Enseignement Supérieur and by the Génopole of Lille. The mass spectrometry facility used in this study was funded by the European Community (FEDER), the Région Nord-Pas de Calais (France), the CNRS and the Université des Sciences et Technologies de Lille.

REFERENCES

1 Allen, A. (1983) Mucus – a protective secretion of complexity. Trends Biochem. Sci. **8**, 169–173
2 Neutra, M. R. and Forstner, J. F. (1987) Gastrointestinal mucus: synthesis, secretion and function. In Physiology of the Gastrointestinal Tract, 2nd edn. (Johnson, L. R., ed.), pp. 975–1009, Raven Press, New York
3 Strous, G. J. and Dekker, J. (1992) Mucin-type glycoprotein. J. Crit. Rev. Biochem. Mol. Biol. **27**, 57–92
4 Forstner, J. F. and Forstner, G. G. (1994) Gastrointestinal mucus. In Physiology of the Gastrointestinal Tract (Johnson, L. R., ed.), pp. 1245–1283, Raven Press, New York
5 Gum, J. R. (1995) Human mucin glycoproteins: varied structures predict diverse properties and specific functions. Biochem. Soc. Trans. **23**, 795–799
6 Van Klinken, B. J., Dekker, J., Buller, H. A. and Einerhand, A. W. (1995) Mucin gene structure and expression: protection vs. adhesion. Am. J. Physiol. **269**, G613–G627
7 Rosen, S. D. and Bertozzi, C. R. (1994) The selectins and their ligands. Curr. Opin. Cell Biol. **6**, 663–673
8 Bresalier, R. S., Ho, S. B., Schoeppner, H. L., Kim, Y. S., Sleisenger, M. H., Brodt, P. and Byrd, J. C. (1996) Enhanced sialylation of mucin-associated carbohydrate structures in human colon cancer metastasis. Gastroenterology **110**, 1354–1367
9 Kim, Y. S. (1998) Mucin glycoproteins in colonic neoplasia. Keio J. Med. **47**, 10–18
10 Brockhausen, I. (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim. Biophys. Acta **1473**, 67–95
11 Kurosaka, A., Nakajima, H., Funakoshi, Z., Matsuyama, M., Nagaya, T. and Yamashina, I. (1983) Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein. J. Biol. Chem. **258**, 11594–11598
12 Capon, C., Wieruszeski, J.-M., Lemoine, J., Byrd, J. C., Leffler, H. and Kim, Y. S. (1997) Sulphated Lewis X determinants as a major structural motif in glycans from LS174T-HM7 human colon carcinoma mucin. J. Biol. Chem. **272**, 31957–31968
13 Podolsky, D. K. (1985) Oligosaccharide structures of human colonic mucin. J. Biol. Chem. **260**, 8262–8271
14 Podolsky, D. K. (1985) Oligosaccharide structures of isolated human colonic mucin species. J. Biol. Chem. **260**, 15510–15515
15 Capon, C., Maes, E., Michalski, J. C., Leffler, H. and Kim, Y. S. (2001) Sd^a-antigen-like structures carried on core 3 are prominent features of glycans from the mucin of normal human descending colon. Biochem. J. **358**, 657–664
16 Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J. P. and Michalski, J. C. (2003) Evidence of regio-specific glycosylation in human intestinal mucins. J. Biol. Chem. **278**, 46337–46348
17 Sangadala, S., Bhat, U. R. and Mendicino, J. (1992) Quantitation and structures of oligosaccharide chains in human trachea mucin glycoproteins. Mol. Cell. Biochem. **118**, 75–90
18 Klein, A., Carnoy, C., Lamblin, G., Roussel, P., van Kuik, J. A. and Vliegenthart, J. F. G. (1993) Isolation and structural characterization of novel sialylated oligosaccharide-alditols from respiratory-mucus glycoproteins of a patient suffering from bronchiectasis. Eur. J. Biochem. **211**, 491–500
19 Chance, D. L. and Mawhinney, T. P. (1996) Disulphated oligosaccharides derived from tracheobronchial mucous glycoproteins of a patient suffering from cystic fibrosis. Carbohydr. Res. **295**, 157–177
20 Robbe, C., Capon, C., Coddeville, B. and Michalski, J. C. (2004) Diagnostic ions for the rapid analysis by nano-electrospray ionization quadrupole time-of-flight mass spectrometry of O-glycans from human mucins. Rapid Commun. Mass Spectrom. **18**, 412–420
21 Zweibaum, A., Oriol, R., Dausset, J., Marcelli-Barge, A., Ropartz, C. and Lanset, S. (1975) Definition in Man of a polymorphic system of the normal colonic secretions. Tissue Antigens **6**, 121–128
22 Lesuffleur, T., Roche, F., Hill, A. S., Lacasa, M., Fox, M., Swallow, D. M., Zweibaum, A. and Real, F. X. (1995) Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cells. The 3′ end of MUC5AC? J. Biol. Chem. **270**, 13655–13673
23 Carlson, D. M. (1968) Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins. J. Biol. Chem. **243**, 616–626
24 Domon, B. and Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj. J. **5**, 397–409
25 Karlsson, N. G., Karlsson, H. and Hansson, G. C. (1996) Sulphated mucin oligosaccharides from porcine small intestine analysed by four-sector tandem mass spectrometry. J. Mass Spectrom. **31**, 560–572
26 Breg, J., van Halbeek, H., Vliegenthart, J. F., Lamblin, G., Houvenaghel, M. C. and Roussel, P. (1987) Structure of sialyl-oligosaccharides isolated from bronchial mucus glycoproteins of patients (blood group O) suffering from cystic fibrosis. Eur. J. Biochem. **168**, 57–68
27 Savage, A. V., Donoghue, C. M., D'Arcy, S. M., Koeleman, C. A. and van den Eijnden, D. H. (1990) Structural determination of sialylated tetrasaccharides and pentasaccharides with blood group H and Le^x activity isolated from bovine submaxillary mucin. Eur. J. Biochem. **192**, 427–432
28 Capon, C., Leroy, Y., Wieruszeski, J. M., Ricart, G., Strecker, G., Montreuil, J. and Fournet, B. (1989) Structures of O-glycosidically linked oligosaccharides isolated from human meconium glycoproteins. Eur. J. Biochem. **182**, 139–152
29 Chai, W., Piskarev, V. and Lawson, A. M. (2001) Negative-ion electrospray mass spectrometry of neutral underivatized oligosaccharides. Anal. Chem. **73**, 651–657
30 Audie, J. P., Janin, A., Porchet, N., Copin, C., Gosselin, B. and Aubert, J. P. (1993) Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization. J. Histochem. Cytochem. **41**, 1479–1485
31 Chang, S.-K., Dohrma, A. F., Basbaum, C. B., Ho, S. B., Tsuda, T., Toribara, N. W., Gum, J. R. and Kim, Y. S. (1994) Localization of mucin (MUC2 and MUC3) messenger RNA and peptide expression in human normal intestine and colon cancer. Gastroenterology **107**, 28–36
32 De Bolos, C., Garrido, M. and Real, F. X. (1995) MUC6 apomucin shows a distinct normal tissue distribution that correlates with Lewis antigen expression in the human stomach. Gastroenterology **109**, 723–734
33 Porchet, N., Nguyen, V. C., Dufosse, J., Audie, J. P., Guyonnet-Duperat, V., Gross, M. S., Denis, C., Degand, P., Bernheim, A. and Aubert, J. P. (1991) Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. Biochem. Biophys. Res. Commun. **175**, 414–422
34 Kurosaka, A., Nakajima, H., Funakoshi, I., Matsuyama, M., Nagayo, T. and Yamashina, I. (1983) Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein. J. Biol. Chem. **258**, 11594–11598
35 Feeney, J., Frenkiel, T. A. and Hounsell, E. F. (1986) Complete 1H-NMR assignments for two core-region oligosaccharides of human meconium glycoproteins, using 1D and 2D methods at 500 MHZ. Carbohydr. Res. **152**, 63–72
36 Montiel, M. D., Krzewinski-Recchi, M. A., Delannoy, P. and Harduin-Lepers, A. (2003) Molecular cloning, gene organization and expression of the human UDP-GalNAc:Neu5Acα2-3Galβ-Rβ1,4-N-acetylgalactosaminyltransferase responsible for the biosynthesis of the blood group Sd^a/Cad antigen: evidence for an unusual extended cytoplasmic domain. Biochem J. **373**, 369–379

© 2004 Biochemical Society

37 Chandrasekaran, E. V., Jain, R. K. and Matta, K. L. (1992) Ovarian cancer α1,3-L-fucosyltransferase: differentiation of distinct catalytic species with the unique substrate, 3′-sulfo-N-acetyllactosamine in conjunction with other synthetic acceptors. J. Biol. Chem. **267**, 23806–23814
38 Smith, A. C. and Podolsky, D. K. (1986) Colonic mucin glycoproteins in health and disease. Clin. Gastroenterol. **15**, 815–837
39 Corfield, A. P., Williams, A. J. K., Clamp, J. R., Wagner, S. A. and Mountford, R. A. (1988) Degradation by bacterial enzymes of colonic mucus from normal subjects and patients with inflammatory bowel disease: the role of sialic acid metabolism and the detection of a novel O-acetylsialic acid esterase. Clin. Sci. **74**, 71–78
40 Corfield, A. P., Wagner, S. A., O'Donnell, L. J. D., Dutdey, P., Mountford, R. A. and Clamp, J. R. (1993) The roles of enteric bacterial sialidase, sialate *O*-acetylesterase and glycosulfatase in the degradation of human colonic mucin. Glycoconj. J. **10**, 72–81
41 Karlsson, K. A. (1989) Animal glycosphingolipids as membrane attachment sites for bacteria. Annu. Rev. Biochem. **58**, 309–350
42 Mikinu-Takagaki, Y. and Hotta, K. (1979) Characterization of peptic inhibitory activity associated with sulphated glycoprotein isolated from gastric mucosa. Biochim. Biophys. Acta **584**, 288–297
43 Roberton, A. M. and Wright, D. P. (1997) Bacterial glycosulphatases and sulphomucin degradation. Can. J. Gastroenterol. **11**, 361–366
44 Nadzielijko, C. and Finkelstein, I. (1994) Inhibition of neutrophil elastase by mucus glycoprotein. Am. J. Resp. Cell Mol. Biol. **11**, 103–107
45 Corfield, A. P., Wagner, S. A., Clamp, J. R., Kriaris, M. S. and Hoskins, L. C. (1992) Mucin degradation in the human colon: production of sialidase, sialate *O*-acetylesterase, *N*-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. Infect. Immun. **60**, 3971–3978
46 Kim, Y. S., Gum, J. R. and Brockhausen, I. (1996) Mucin glycoproteins in neoplasia. Glycoconj. J. **13**, 693–707

Received 13 April 2004/6 September 2004; accepted 13 September 2004
Published as BJ Immediate Publication 13 September 2004, DOI 10.1042/BJ20040605

© 2004 Biochemical Society